



Copy of RECEIPT  
#6  
PCT  
2/21

**PATENT  
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

IN RE APPLICATION: Usama Fayyad et al.

**RECEIVED**

Serial No. 09/700,606

SEP 06 2001

Filed: January 31, 2001

Technology Center 2100

Title: SCALABLE SYSTEM FOR CLUSTERING OF LARGE DATABASES HAVING MIXED ATTRIBUTES

Docket No. 14-694 - 4254

WATTS, HOFFMANN, FISHER  
& HEINKE, CO., L.P.A.  
1100 Superior Avenue, Ste. 1750  
Cleveland, Ohio 44114

Assistant Commissioner for Patents  
Office of Initial Patent Examination's  
Customer Service Center  
Washington, D.C. 20231

**REQUEST FOR CORRECTED FILING RECEIPT**

Sir:

As evidenced by the enclosed copy of page one of the specification, specifically, the first two paragraphs which have been highlighted for your convenience, and the enclosed copy of the Declaration and Power of Attorney, the correct domestic priority data as claimed by applicant should read as follows:

THIS APPLICATION IS A 371 OF PCT/US99/06717 03/29/1999

WHICH IS A CIP OF 09/083,906 05/22/1998

WHICH CLAIMS BENEFIT OF 60/086,410 05/22/1998

I hereby certify that this paper is being deposited with the U.S. Postal Service as 1st Class Mail addressed to the Assistant Commissioner of Patents, Washington, D.C. 20231

on

By

6/6/01  
Maren Burroughs

**BEST AVAILABLE COPY**

A copy of the original filing receipt is enclosed which shows the necessary corrections. A corrected filing receipt is respectfully requested.

Respectfully submitted,

Date: June 1, 2001

Stephen S. Schultz  
Stephen S. Schultz  
Reg. No. 29,108

Telephone: 216-241-6700  
Facsimile: 216-241-8151



## UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS  
UNITED STATES PATENT AND TRADEMARK OFFICE  
WASHINGTON, D.C. 20231  
www.uspto.gov

APPLICATION NUMBER	FILING DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	DRAWINGS	TOT CLAIMS	IND CLAIMS
09/700,606	01/31/2001	2121	1152	14-694-4254	8	34	4

CONFIRMATION NO. 8142

## FILING RECEIPT



\*OC000000005992681\*

Watts Hoffman Fisher & Heinke Company  
Suite 1750  
1100 Superior Avenue  
Cleveland, OH 44114-2518

APR 27 2001

Date Mailed: 04/24/2001

Receipt is acknowledged of this nonprovisional Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to the Office of Initial Patent Examination's Customer Service Center. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the PTO processes the reply to the Notice, the PTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).

## Applicant(s)

Usama Fayyad, Mercer Island, WA;  
Paul S. Bradley, Seattle, WA;  
Cory A. Reina, Kirkland, WA;

RECEIVED

SEP 06 2001

Technology Center 2100

## Domestic Priority data as claimed by applicant

THIS APPLICATION IS A 371 OF PCT/US99/06717 03/29/1999  
WHICH IS A CIP OF 60/083,906 05/01/1998  
WHICH CLAIMS BENEFIT OF 60/086,410 05/22/1998

09/083,906

05/22/98

## Foreign Applications

Projected Publication Date: N/A

Non-Publication Request: No

Early Publication Request: No

## Title

Scalable system for clustering of large databases having mixed data attributes

Preliminary Class

700

---

Data entry by : NGUYEN, SON

Team : OIPE

Date: 04/24/2001





**LICENSE FOR FOREIGN FILING UNDER  
Title 35, United States Code, Section 184  
Title 37, Code of Federal Regulations, 5.11 & 5.15**

**RECEIVED**

SEP 06 2001

Technology Center 2100

**GRANTED**

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 36 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Office of Export Administration, Department of Commerce (15 CFR 370.10 (j)); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

**NOT GRANTED**

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15 (b).

**PLEASE NOTE the following information about the Filing Receipt:**

- The articles such as "a," "an" and "the" are not included as the first words in the title of an application. They are considered to be unnecessary to the understanding of the title.
- The words "new," "improved," "improvements in" or "relating to" are not included as first words in the title of an application because a patent application, by nature, is a new idea or improvement.
- The title may be truncated if it consists of more than 600 characters (letters and spaces combined).
- The docket number allows a maximum of 25 characters.
- If your application was submitted under 37 CFR 1.10, your filing date should be the "date in" found on the Express Mail label. If there is a discrepancy, you should submit a request for a corrected Filing Receipt along with a copy of the Express Mail label showing the "date in."
- The title is recorded in sentence case.

Any corrections that may need to be done to your Filing Receipt should be directed to:

Assistant Commissioner for Patents  
Office of Initial Patent Examination  
Customer Service Center  
Washington, DC 20231



## A Scalable System for Clustering of Large Databases Having Mixed Data Attributes

### Cross Reference to Related Applications

5 The present invention is a continuation in part of co-pending United States patent application serial number 09/083,906 entitled "A Scalable System for Expectation Maximization Clustering of Large Databases" to Fayyad et al having a filing date of May 22, 1998, and which is assigned to the assignee of the present application and is incorporated herein by reference.

10 Priority is claimed from provisional United States patent application serial number 60/086,410 filed May 22, 1998.

RECEIVED

SEP 06 2001

### Field of the Invention

Technology Center 2100

15 The present invention concerns database analysis and more particularly concerns an apparatus and method for clustering of data into groups that capture important regularities and characteristics of the data.

### Background Art

20 Large data sets are now commonly used in most business organizations. In fact, so much data has been gathered that asking even a simple question about the data has become a challenge. The modern information revolution is creating huge data stores which, instead of offering increased productivity and new opportunities, can overwhelm the users with a flood of information. Tapping into large databases for even simple browsing can result in a return of irrelevant and unimportant facts. Even people who do not 'own' large databases face the overload problem when accessing  
25 databases on the Internet. A large challenge now facing the database community is how to sift through these databases to find useful information.

Existing database management systems (DBMS) perform the steps of reliably storing data and retrieving the data using a data access language, typically SQL. One major use of database technology is to help individuals and organizations make  
30 decisions and generate reports based on the data contained in the database.

An important class of problems in the areas of *decision support* and *reporting* are clustering (segmentation) problems where one is interested in finding groupings (clusters) in the data. Data clustering has been used in statistics, pattern recognition,

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☒ **SKewed/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**